Behavioral/Systems/Cognitive

Action and Outcome Encoding in the Primate Caudate Nucleus

Brian Lau and Paul W. Glimcher

Center for Neural Science, New York University, New York, New York 10003

The basal ganglia appear to have a central role in reinforcement learning. Previous experiments, focusing on activity preceding movement execution, support the idea that dorsal striatal neurons bias action selection according to the expected values of actions. However, many phasically active striatal neurons respond at a time too late to initiate or select movements. Given the data suggesting a role for the basal ganglia in reinforcement learning, postmovement activity may therefore reflect evaluative processing important for learning the values of actions. To better understand these postmovement neurons, we determined whether individual striatal neurons encode information about saccade direction, whether a reward had been received, or both. We recorded from phasically active neurons in the caudate nucleus while monkeys performed a probabilistically rewarded delayed saccade task. Many neurons exhibited peak responses after saccade execution (77 of 149) that were often tuned for the direction of the preceding saccade (61 of 77). Of those neurons responding during the reward epoch, one subset showed direction tuning for the immediately preceding saccade (43 of 60), whereas another subset responded differentially on rewarded versus unrewarded trials (35 of 60). We found that there was relatively little overlap of these properties in individual neurons. The encoding of action and outcome was performed by largely separate populations of caudate neurons that were active after movement execution. Thus, striatal neurons active primarily after a movement appear to be segregated into two distinct groups that provide complimentary information about the outcomes of actions.

Key words: basal ganglia; oculomotor; reward; reinforcement learning; monkey; caudate

Introduction

Learning the values of actions through experience proceeds in two phases. An action is selected based (at least partially) on the expected values of the actions available to the subject. After an action is completed, the outcome of the movement can be used to revise these expectations. There is evidence that the basal ganglia are involved in both of these phases of reinforcement learning (Houk et al., 1995a; Schultz et al., 1995; Schultz, 1998). Individual phasically active neurons (PANs) in the striatum respond to specific events during instrumental conditioning tasks, and as a population, these neurons are responsive from well before action selection to well after reward receipt (Schultz et al., 1995; Schmitzer-Torbert and Redish, 2004; Barnes et al., 2005). Given the available data, it thus seems possible that premovement activity participates in action selection, whereas postmovement activity is important for updating expectations.

A role for PANs in action selection is supported by physiological experiments in primates showing that some neurons with premovement activity can encode reward expectations in instrumental conditioning tasks (Hikosaka et al., 1989c; Apicella et al.,

1992; Hollerman et al., 1998). Neurons active before movement also show learning-related changes when task contingencies are changed (Kawagoe et al., 1998; Tremblay et al., 1998; Lauwereyns et al., 2002a; Takikawa et al., 2002; Pasupathy and Miller, 2005; Watanabe and Hikosaka, 2005), and respond in a manner consistent with a causal role in biasing the speed (Lauwereyns et al., 2002b; Itoh et al., 2003; Watanabe et al., 2003) or selection (Samejima et al., 2005) of particular actions.

Although these observations provide compelling evidence that striatal PANs encode the learned values of actions and may influence the speed and frequency with which actions are selected, there are also reports of striatal neurons responding after movement execution and after reward delivery (Rolls et al., 1983; Hikosaka et al., 1989c; Apicella et al., 1991). These postmovement responses have not been studied extensively, and although it is clear that the timing of these responses is inappropriate for initiating movements, how these neurons would fit into models of the basal ganglia remains unclear. Previous reports have shown that postmovement activity in the dorsal striatum can depend on response modality and task requirements, as well as on learning (Hikosaka et al., 1989c; Apicella et al., 1991; Hollerman et al., 1998; Tremblay et al., 1998; Williams and Eskandar, 2006). Two questions, however, remain unanswered. First, do postmovement neurons encode both reward properties and movement metrics? Second, how do they encode these properties, if indeed they encode them at all? Individual PANs, for example, might independently encode both the direction of a movement and the outcome associated with that action. If this were the case, then the

fact that a neuron encodes movement direction would not predict whether it also encodes the obtained outcome. Alternatively, individual PANs might interactively encode both the direction of a movement and the outcome associated with that action; if this were the case, the fact that a neuron encodes movement direction might well predict whether or not it also encodes the obtained outcome.

To better understand postmovement neurons in the striatum, we studied the activity of neurons in the caudate nucleus while monkeys performed a saccade task for probabilistic rewards, which allowed us to measure postsaccadic neuronal responses on both rewarded and unrewarded trials. We varied the direction of the instructed saccades and characterized the direction tuning of neurons throughout the task. We held reward expectation constant by fixing the probability of reward. We found that a subset of postmovement neurons were sharply tuned for the direction of the preceding saccade, whereas a different subset of postmovement neurons were differentially responsive to rewarded versus unrewarded trials. These two PAN populations may play different roles in evaluative updating, one providing information about the saccade selected and the other encoding the outcome received.

Materials and Methods

Subjects and surgery. Two male rhesus monkeys (*Macaca mulatta*) were used as subjects (monkey B and monkey H, ~10.5 kg and ~11.5 kg, respectively). All experimental procedures were designed in association with the university veterinarian, approved by the New York University Institutional Animal Care and Use Committee, and performed in compliance with the Public Health Service's *Guide for the Care and Use of Animals.*

Before behavioral training, each animal was implanted with a head restraint prosthesis and a scleral eye coil (Judge et al., 1980) to allow for the maintenance of stable head position and recording of eye position. Surgical procedures were performed using conventional aseptic techniques under general anesthesia (detailed in Platt and Glimcher, 1997). Analgesia and antibiotics were administered during surgery and continued for 3 d postoperatively. Training began after a postoperative recovery period of 6 weeks.

After the monkeys had been trained to execute the oculomotor task used in this study, a second sterile surgical procedure was performed to implant a circular recording chamber (2 cm diameter, CILUX plastic for monkey B, stainless steel for monkey H; Crist Instruments, Hagerstown, MD). The recording chamber was centered stereotaxically over the body of the caudate nucleus (3 mm posterior to the anterior commissure), 5 mm lateral to the midline, and oriented perpendicular to the stereotaxic horizontal plane. A portion of the underlying skull was removed (15 mm diameter), and the recording chamber was bonded to four additional orthopedic bone screws (Synthes, West Chester, PA) and the original implant with methyl methacrylate acrylic cement (Biomet, Warsaw, IN). The recording chamber was regularly cleaned with sterile saline, and antibiotics were used as necessary.

Experimental setup. Experiments were conducted in a dimly lit sound-attenuated room. The monkeys were head restrained and seated in a Plexiglas enclosure (28 × 48 cm). Body movements were monitored from a separate room using an infrared camera.

Eye movements were measured using the scleral search coil technique (Fuchs and Robinson, 1966). Horizontal and vertical eye positions were detected and calibrated using a quadrature phase detector (Riverbend Electronics, Birmingham, AL) and sampled at 500 Hz.

Visual stimuli were generated using an array of light-emitting diodes (LEDs) positioned 145 cm from the monkeys' eyes. The array contained 567 LEDs (21 × 27, 2° spacing), with each LED subtending ~0.25° of visual angle.

Behavioral task. The data reported in this study were collected while the monkeys performed a delayed saccade task (Fig. 1). Each trial started with a 300 ms 500 Hz tone, after which the monkey was given 700 ms to
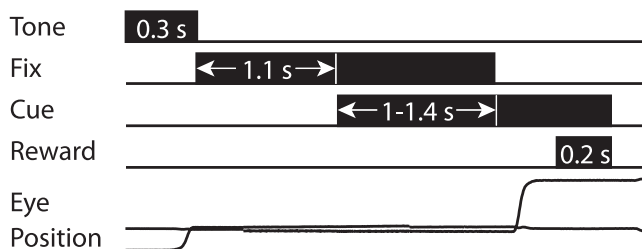

Figure 1. Timeline of delayed saccade task. The fixation point and the peripheral cue were coilluminated during the delay, which was randomly selected to be 1, 1.2, or 1.4 s in duration.

align its gaze to within 2–3° of a yellow LED in the center of the visual field. After maintaining fixation for 400 ms, a peripheral red LED was illuminated at one of eight possible locations arranged symmetrically around the fixation point. These peripheral LEDs were positioned an equal distance from the fixation point (average eccentricity ~16° with a ~2° variation resulting from rectilinear LED spacing). The central fixation point was extinguished after a random delay of 1000, 1200, or 1400 ms, which cued the monkey to shift gaze within 450 ms to the peripheral LED (±3–4°). Note that the fixation point and the peripheral cue were coilluminated for the duration of the delay; this task has also been referred to as an overlap saccade task (Hikosaka et al., 1989a). Rewards were delivered randomly at a fixed percentage of 30–50% of completed trials. We occasionally varied this percentage within a session, but it was held constant for the duration of data collection from any particular neuron. If a reward was scheduled, 0.38 ml of water was delivered through a sipper-tube after the monkey fixated for 200 ms after acquisition of the peripheral cue. Trial timing remained the same whether or not the trial had been randomly selected for reward delivery, and the monkey was required to maintain fixation for the duration of the reward epoch (200 ms), regardless of whether a reward was delivered, in order for the trial to be recorded as correctly completed. All correctly completed trials were accompanied by a 200 ms noise burst that served as a secondary reinforcer. Trials were separated by a 1500–2500 ms intertrial interval.

We considered a trial aborted if the monkey failed to align its gaze within the required distance of the fixation or cue LEDs, if an eye movement was made to the cue before the fixation point was extinguished, or if fixation at the peripheral LED was broken before the completion of the reward epoch (<1% of aborts). When an abort was detected (most of which were failures to acquire fixation or broken fixations before the peripheral cue was even presented), any illuminated LEDs were extinguished immediately, neither water nor the secondary reinforcer was delivered, and the next trial began after a 3000–6000 ms time-out.

Electrophysiological recording. At the start of each recording session, an *x–y* micropositioner (Crist Instruments) and a hydraulic microdrive (Kopf Instruments, Tujunga, CA) were mounted to the recording chamber, into which a guide tube support grid had been secured (1 mm spacing; Crist Instruments). A 23 gauge sharpened guide tube, housing a tungsten steel electrode (2–4 MΩ measured at 1 kHz, 0.25 mm diameter shaft; FHC, Bowdoin, ME), was used to puncture the dura. The guide tube was lowered until its tip was above or just lateral to the cingulate sulcus as predetermined using magnetic resonance imaging (3T; Siemens, Ehrlangen, Germany) in monkey B and B-mode ultrasound imaging (GE Healthcare, Little Chalfont, UK) (Glimcher et al., 2001) in monkey H. Electrophysiological signals were amplified, bandpass filtered (500–10,000 Hz), and displayed on an oscilloscope. Individual neurons were isolated using a time–amplitude window discriminator with an adjustable Schmitt trigger (Bak Electronics, Germantown, MD), which produced transistor–transistor logic pulses corresponding to the times of single action potentials. Spike times were recorded with 1 µs resolution.

We lowered electrodes under physiological guidance until we encountered the dorsal surface of the caudate nucleus. We used a vertical approach, and as the electrode exited the guide tube, we normally encountered activity from the cingulate sulcus. After exiting this area, we encountered several millimeters of white matter, identified by a lack of neuronal activity punctuated occasionally by fiber potentials. When the

electrode reached the caudate nucleus, we normally encountered injury potentials as well as spontaneous activity.

Once a caudate neuron was isolated, we had the monkey perform delayed saccades and simple fixations to assess neural responsiveness. We distinguished putative projection neurons from tonically active neurons (TANs) based on differences in spontaneous activity, spike waveform, response to reward, and movement selectivity (Kimura et al., 1984; Hikosaka et al., 1989a; Aosaki et al., 1994; Apicella et al., 1997; Lee et al., 2006). PANs were characterized as those neurons having low spontaneous activity (1.1 ± 1 spikes/s, mean $\pm$ SD), whereas TANs were identified as those neurons having a higher rate of spontaneous activity (5.3 ± 1.2 spikes/s), broader spike waveform, characteristic response to reward, and lack of clear movement selectivity. When we isolated a PAN and judged the neuron to be responsive in the task (typically by observing a phasic response to one or more task events), we had the monkey perform a series of delayed saccade trials to eight possible cue directions (45° angular spacing, eccentricity ~16°) to determine direction tuning and responsiveness to obtained outcome. We recorded from each neuron for an average of 123 correctly completed trials ($\pm$51, SD), for an average of 15 trials for each direction (minimum three trials per direction).

Histology. The location of the recording sites was verified histologically in monkey H. Lesions were made at the site of neuronal recordings by passing a 5 μA anodal current through the recording electrode for 10 s. The animal was premedicated with ketamine and then killed with an overdose of sodium thiopental. It was perfused intracardially with a saline solution followed by 4% paraformaldehyde in PBS and finally by 30% sucrose in PBS. The brain was then removed from the skull, submerged in 30% sucrose for several days, and then blocked and cut into 40 μm frozen sections. The sections were mounted and stained for Nissl substance, and camera lucida drawings were made of selected sections. We constructed a dorsal view of the caudate showing the approximate positions of our recording sites by reconstructing sections aligned to guide pins inserted at specific locations within the recording chamber.

Data analysis. We characterized three properties for each neuron: (1) tuning sharpness, the degree to which a neuron responded more strongly for some directions versus others; (2) preferred direction; and (3) reward responsiveness, the degree to which a neuron responded differentially to rewarded versus unrewarded trials. We used circular statistics to measure these properties (Fisher, 1993).

We characterized tuning sharpness using the magnitude of the vector sum of the eight cue directions, weighted by the corresponding firing rate r_k, for each direction θ_k, and normalized by the overall response:

$$TS = \left| \frac{\sum_{k=1}^{8} r_k e^{i\theta_k}}{\sum_{k=1}^{8} r_k} \right|.$$

This statistic (also known as the mean resultant length = $1 -$ circular variance) assumes a value of 0 if a neuron responds equally to all directions (no tuning) and a value of 1 if a neuron responds to only one direction (sharpest tuning). The preferred direction (*PD*) was estimated as the circular mean (vector mean direction). We use the convention that 0° is movement along the horizontal meridian contralateral to the recorded hemisphere.

We characterized reward responsiveness by first estimating a response strength separately for rewarded and unrewarded trials for all eight cue directions. Response strength was measured using the radius of a circle with an area equal to the area of each tuning curve: $M = (A/\pi)^{1/2}$, where A is the area of the tuning curve estimated using the method developed by Gribble and Scott (2002). We used response strengths estimated for re-
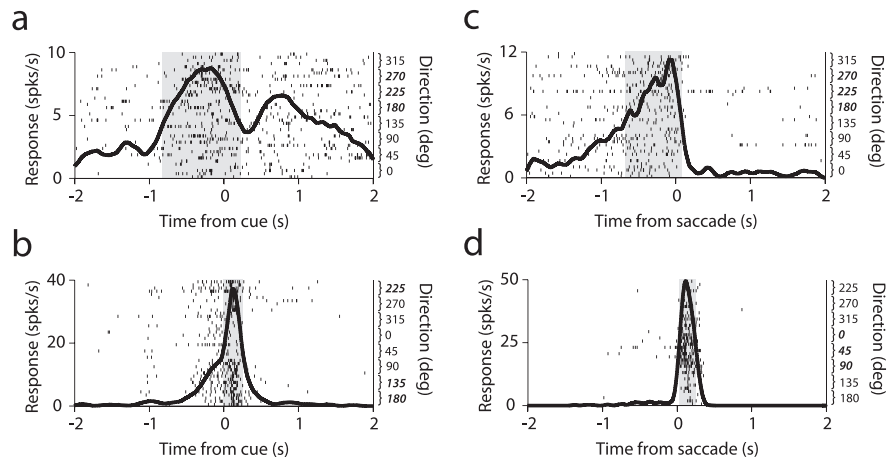
Figure 2. Temporal response profiles for four example PANs. *a*, Neuron responding before and during cue presentation (Gaussian smoothing $\sigma = 100$ ms). Rasters show five trials for each direction arranged according to the right axis (only a subset of the total trials recorded is plotted for clarity). Spike density functions were estimated from the three cue directions (italicized directions on the right axis) that elicited the largest responses. The gray regions represent the first times to half-maximal response before and after the peak response time (see Materials and Methods). *b*, Neuron responding to cue presentation ($\sigma = 55$ ms). *c*, Neuron responding before saccade ($\sigma = 45$ ms). *d*, Neuron responding immediately after the completion of a saccade, but before the onset of reward ($\sigma = 32$ ms).

warded (M_R) and unrewarded trials (M_U) to calculate a reward responsiveness index: $RR = (M_R - M_U)/[\max(M_R, M_U)]$, which yields an estimate of the difference in response strength between rewarded and unrewarded trials normalized to allow for comparison across neurons. This index provides a measure of reward responsiveness that is independent of direction tuning. It assumes a value of 0 if the neuronal response is unaffected by whether or not the trial was rewarded ($M_R = M_U$). It assumes values of -1 or $+1$ when a neuron responds exclusively during unrewarded or rewarded trials, respectively.

We determined the statistical significance for the circular statistics using a parametric bootstrap (Efron and Tibshirani, 1994; Stark and Abeles, 2005). We used bootstrap tests to avoid deriving analytic uncertainty intervals for the measures described above; standard asymptotic tests for circular statistics (e.g., Fisher, 1993) are not applicable to our measurements because we repeatedly sampled a variable (firing rate) at eight fixed directions. For a particular time window for a given neuron, we generated 2000 bootstrap replications of *TS*, *M*, and *RR* by randomly sampling from a Poisson distribution (Hanes et al., 1995) with a mean equal to that of the actual data for each direction. We chose not to resample from the empirical distribution of firing rates because the number of trials per direction was small enough to bias estimation (Efron and Tibshirani, 1994); therefore, our tests are not nonparametric. We computed confidence intervals (CIs, 95%) by sorting the bootstrap replications and taking the values of the 2.5th and 97.5th percentiles as the lower and upper CIs, respectively. We deemed *TS*, *M*, and *RR* significant at the $p < 0.05$ level when the CIs for these did not enclose zero.

In the first portion of this report, we summarize the tuning and reward properties across groups of neurons. Because caudate neurons have variable response profiles (both in the timing of peak responses and in the distribution of activity in time), we used temporal analysis windows tailored for each neuron (Fig. 2). First, we estimated a spike density function using a Gaussian smoothing window, combining all trials from the three cue directions that elicited the largest response over the entire trial. The bandwidth used for the Gaussian smoothing was chosen to maximize the information gain per spike (Paulin and Hoffman, 2001) for each neuron. Next, the peak response for each neuron was estimated as the maximum of the spike density function, and an analysis window was defined using the first times to half-maximum before and after the peak response (half-max windows). This method succeeded in capturing a significant portion of the neuronal response while maintaining good signal-to-noise in our other measures.

Finally, we examined the temporal evolution of tuning sharpness over the course of a trial for single neurons. We did this by computing the
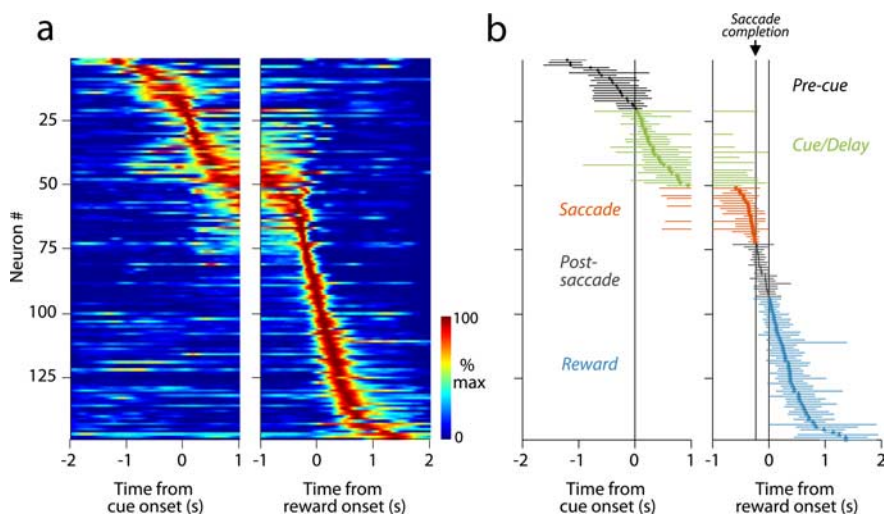
Figure 3. Temporal response profiles of all neurons during the delayed saccade task ($n = 149$). **a**, Each row represents a spike density function for a single neuron, estimated using trials from the three target directions that elicited the greatest response for each neuron. Each spike density function was normalized by its peak response, and the data are sorted by peak response time. **b**, Individual peak response times (symbols) with first times to half-maximal response before and after the peak response time (half-max windows, beginning and end of the line for each row). Neurons are plotted in the same order as in **a**. The different colors represent the neuronal categories defined in the Materials and Methods.

tuning sharpness in 200 ms nonoverlapping windows across the entire trial.

Neuron classification. We classified our neuronal population into five exclusive categories based on the timing of peak responses relative to three task events: cue onset, saccade execution, and reward onset. Precue neurons had peak response times preceding cue onset. Cue/Delay neurons had peak response times after cue onset and preceding saccade onset by at least 300 ms. Saccade neurons had peak response times within a 300 ms window preceding saccade completion. Postsaccade neurons had peak response times after saccade completion, but before reward onset. Reward neurons had peak response times after reward onset. Finally, we refer to neurons in the Precue, Cue/Delay, and Saccade categories as premovement neurons, and those in the Postsaccade and Reward categories as postmovement neurons. Note that we chose to use saccade completion to define our categories to be conservative about including neurons into the Postsaccade category. In practice, all but two Saccade neurons had peak responses that preceded saccade onset. The remaining two neurons had peak responses when the eye was moving, and we include them in the Saccade category.

Results

We recorded from 373 neurons from two caudate nuclei in two monkeys. Two hundred thirty-four neurons were judged to be PANs. The remaining 139 neurons were judged to be tonically active neurons. Of the 234 putative projection neurons, we report here on 149 (86 from monkey B, 63 from monkey H) that showed phasic increases in activity relative to baseline at some point in the delayed saccade task (Fig. 1) and were held for a sufficient number of trials.

These PANs exhibited a variety of response profiles. Example neurons are shown in Figure 2; the first two examples responded around the time of cue presentation, whereas the second two examples responded around the time of the saccadic eye movement. These examples also illustrate the response timescales we observed. These differences in response timing and timescale indicated that using a single temporal analysis window for the entire population would be inappropriate. Instead, for each neuron, we used the first times before and after the peak response time to define an analysis window (half-max window). In Figure 2, the duration of the half-max window is ~1000 ms for the first exam-

ple, and ~250 ms for the last example (represented by the gray shading in each panel). Despite these differences in duration, the half-max windows captured a significant portion of the response for each neuron, as might be expected based on the characteristic responses of projection neurons in the oculomotor portion of the caudate nucleus (e.g., Hikosaka et al., 1989a,b,c).

Figure 3*a* presents the average response profile for each neuron, normalized by peak response and rank sorted by peak response time. The timing of peak responses across our population spanned the entirety of the trial, from well before cue presentation to well after saccade execution and into the reward epoch. In fact, 52% (77 of 149) of the neurons had peak responses that followed saccade completion. The peak response times and half-max windows are plotted for each neuron in Figure 3*b*. As was the case for the examples in Figure 2, the individual half-max windows captured a significant portion of the response for each neuron across our population. We therefore focus the first half of this report on analyses using these windows, deferring a dynamic analysis examining activity throughout the trial until the penultimate section of this report.

Figure 3*b* also illustrates our classification of each neuron based on the timing of peak responses relative to cue onset, saccade execution, and reward onset (see Materials and Methods): Precue (13%), Cue/Delay (20%), Saccade (15%), Postsaccade (14%), and Reward (38%). The proportions of neurons assigned to each category were not significantly different between monkeys for any category ($p > 0.10$, χ^2).

Neurons with peak responses preceding reward onset

In this section, we characterize the direction tuning of PANs with peak responses preceding reward onset ($n = 93$; 20 Precue, 30 Cue/Delay, 22 Saccade, 21 Postsaccade). The direction tuning for eight representative neurons is plotted in Figure 4 (two neurons from each category except Reward, examples of which are presented in the next section). PANs displayed a wide range of tuning profiles, ranging from untuned to sharply tuned. These neurons also exhibited preferred directions that included contralateral as well as ipsilateral cue and saccade directions. A subset of neurons responding well after saccade execution showed clear direction tuning for the immediately preceding saccade (e.g., the Postsaccade neuron in Fig. 4, top right).

We used circular statistics to characterize the TS and the PD for each neuron (see Materials and Methods for details). We found that most of the neurons in the Cue/Delay, Saccade, and Postsaccade categories were significantly tuned ($p < 0.05$, bootstrap test for TS) to certain cue or saccade directions, although tuning sharpness was broadly distributed (Fig. 5*a*). There was a significant difference between the average significant TS of the Cue/Delay, Saccade, and Postsaccade neurons ($p < 0.05$, ANOVA), and *post hoc* tests revealed that Saccade neurons were more sharply tuned than Postsaccade neurons ($p < 0.05$, Tukey–Cramer). On the other hand, none of the Precue neurons were significantly tuned for cue direction. This is expected for those Precue neurons responding entirely before cue onset (7 of 20 =

35%), when the monkey did not yet know which saccade would be instructed. However, tuning was also not observed for the 65% (13 of 20) of Precue neurons with half-max windows extending into the cue epoch. The *TS* values estimated for Precue neurons provide an independent check on the validity of our statistical test for tuning strength. Most significantly direction-tuned neurons had *TS* values greater than those observed for Precue neurons, although a small number of neurons exhibited statistically reliable, but weak, direction tuning (in the range observed for Precue neurons).

For significantly direction-tuned neurons, preferred directions were significantly biased toward contralateral directions ($p < 0.01$, Rayleigh test), with a circular mean of 352 ± 29° (±95% CI). However, when tested by category, the contralateral bias was significant for the Postsaccade ($p < 0.01$) but not the Cue/Delay ($p = 0.90$) or Saccade ($p = 0.11$) neurons. We summarized tuning across the population using average tuning curves (Fig. 5c).

The direction tuning observed in Cue/Delay and Saccade neurons is related to the visual cue presentation or to the forthcoming saccade, and possibly both. These neurons encode information appropriate for the planning and execution of the instructed saccade. Postsaccade neurons, on the other hand, encode retrospective direction information; tuning is with respect to the direction of saccades already completed.

Figure 4. Phasically active neurons exhibit a variety of tuning for cue or saccade direction. Polar plots of direction tuning for eight example neurons (two neurons each for Precue, Cue/Delay, Saccade, and Postsaccade categories; examples for the Reward category are shown in Figs. 6, 8). The *TS* and response strength (*M*) are also listed for each neuron. Contralateral movements are to the right. The panels marked with asterisks correspond to the neurons in Figure 2. Firing rates were estimated using the times to half-maximal response before and after the peak response time for each neuron. The half-max windows were as follows: −757 to 301 ms and −758 to 246 ms relative to cue onset for the Precue neurons; 166 to 1262 ms and −2 to 287 ms relative to cue onset for the Cue/Delay neurons; −518 to 50 ms and −642 to 73 ms relative to saccade completion for the Saccade neurons; 13 to 238 ms and −53 to 208 ms relative to saccade completion for the Postsaccade neurons.

Neurons active during reward epoch
Rewards were delivered on a random fraction of trials (0.30–0.50), and this fraction was held constant while recording from a neuron; therefore neuronal activity during the reward epoch could potentially distinguish between rewarded and unrewarded trials. All correctly executed saccades yielded a secondary auditory reinforcer, meaning that both rewarded and unrewarded trials were correct if the instructed saccade was executed. Errors were indicated by a lack of secondary reinforcer and a time out; error trials were rare, and were excluded from further analysis. We examined neurons responding during the reward epoch ($n = 60$; 4 Postsaccade, 56 Reward) to determine the following: (1) whether they were responsive to reward outcome, (2) whether they were retrospectively direction tuned, and (3) whether individual neurons jointly encoded action and received outcome. A small subset of Postsaccade neurons responded, albeit less vigorously, during reward delivery and could therefore exhibit significant reward responsiveness. For completeness, we include four Postsaccade neurons in our analyses of reward epoch responses because their half-max windows extended for at least 150 ms of reward delivery (not the case for any Precue, Cue/Delay, or Saccade neurons).

Tuning profiles for representative Reward neurons are shown in Figure 6a, where the data are plotted separately for rewarded and unrewarded trials. As illustrated by the examples, we found that neuron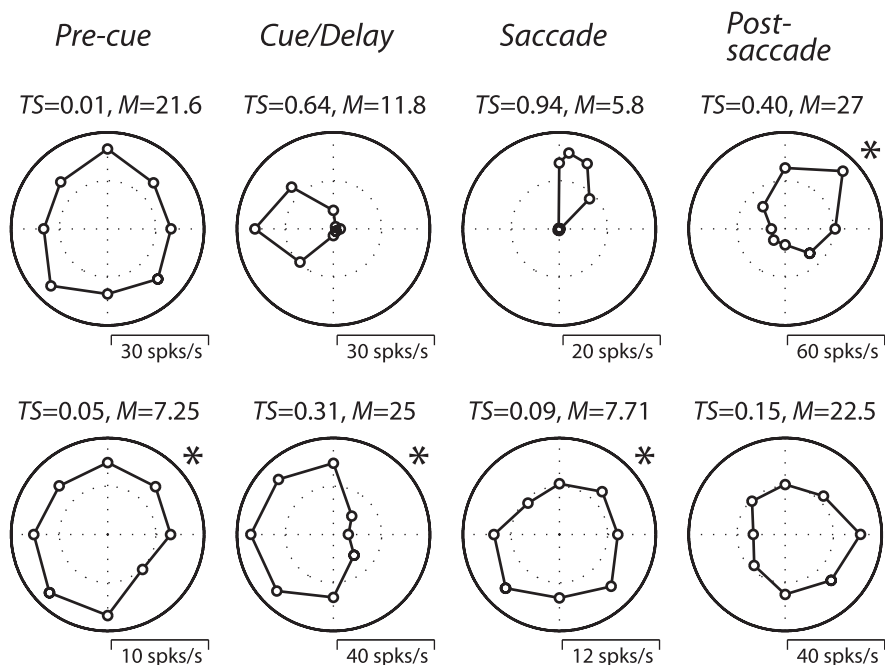s could be differentially responsive for reward, tuned for cue or saccade direction, or somewhat responsive to both reward and direction. We quantified differences between rewarded and unrewarded responses using a reward responsiveness index (*RR*), which measures how much more or less responsive a neuron is on rewarded trials (see Materials and Methods for details). Approximately one-half (35 of 60; 1 Postsaccade, 34 Reward) of the neurons active during the reward epoch were differentially responsive to reward delivery ($p < 0.05$, bootstrap test for *RR*). The majority of these neurons (63%, 22 of 35) had a modulation >50% ($|RR| > 0.5$) in response to reward. We observed neurons with both signs of reward response (Fig. 6a, top two examples); 57% of significant neurons (20 of 35) showed larger response during unrewarded trials relative to rewarded trials, with the remaining 43% (15 of 35) showing the reverse (Fig. 6b).

We used *TS* and *PD* to quantify direction tuning, calculating each statistic separately for rewarded and unrewarded trials (Fig. 7). Seventy-two percent of these neurons (43 of 60) were significantly direction tuned for at least one reward condition ($p < 0.05$, bootstrap test for *TS*). This direction tuning occurred quite late in the trial; statistics for the vast majority of these neurons were computed using half-max windows that began after the saccade to the cue was completed. Neurons showed similar tuning for rewarded and unrewarded trials. The neurons significantly tuned for both reward conditions (51%, 22 of 43) tended to be more sharply tuned; these neurons showed a strong agreement between tuning sharpness (Pearson's linear correlation $r = 0.84$, $p < 0.001$) and preferred direction (T-monotone circular-circular association = 0.88, $p < 0.001$) for rewarded and unrewarded trials. This indicates that some neurons responding during the reward epoch were sharply tuned for saccade direction
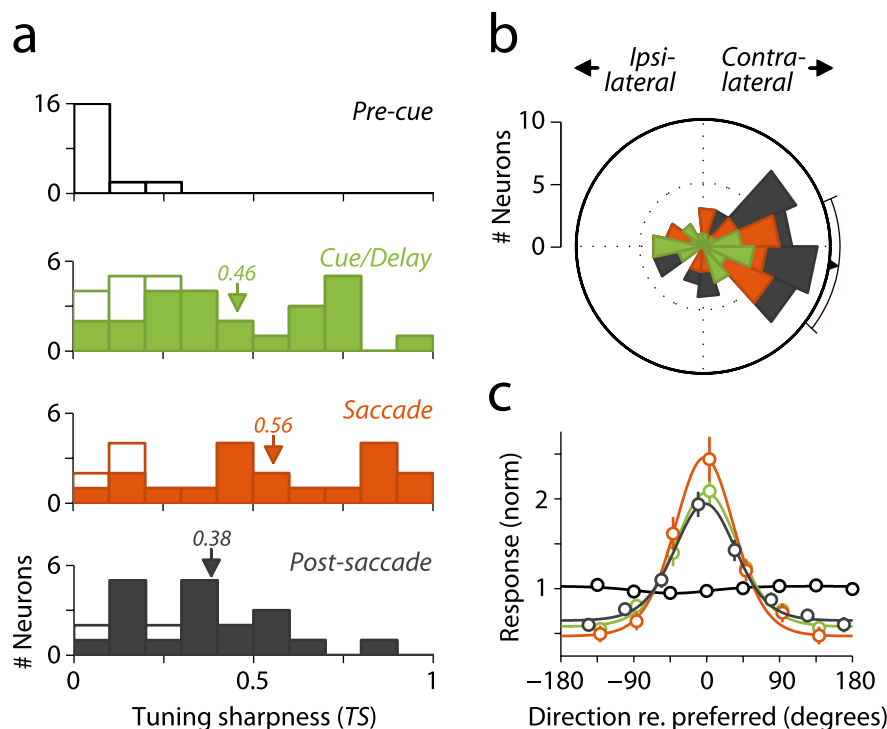
a

b

c

Figure 5. Direction-tuning summary for neurons with peak responses preceding reward onset. **a**, Histograms of *TS*. Colored bars indicate statistically significant direction tuning ($p < 0.05$, bootstrap test for *TS*). Arrows indicate means for significant *TS*. There were no significantly tuned Precue neurons. **b**, Stacked polar histogram of preferred directions for significantly tuned neurons. The population preferred directions are significantly biased toward contralateral directions ($p < 0.01$, Rayleigh test), with a circular mean of 352 ± 29° (± 95% CI, triangle and arc along outer edge). **c**, Population tuning functions. Each category is plotted separately, and except for the Precue category, only significantly tuned neurons are included in the average. The symbols represent the mean (±1 SEM), and the lines represent the best-fitting von Mises (circular Gaussian) function. The firing rate of each neuron was normalized to its average response to all directions; a value of 1 indicates no modulation from the mean (i.e., untuned).

independently of whether or not reward was delivered. Similar to neurons with peak responses preceding reward onset, neurons responding during the reward epoch displayed a broad range of tuning profiles. The average tuning sharpness was 0.35, which indicates that neurons responding during the reward epoch were, on average, slightly less sharply direction tuned than Cue/Delay, Saccade, and Postsaccade neurons.

The preferred directions for significantly tuned neurons had a circular mean of 235 ± 46° (±95% CI). In contrast to the neurons with peak responses preceding reward onset, the preferred directions across the population of neurons responding during the reward epoch were not biased toward contralateral directions in a unimodal manner ($p > 0.10$, Rayleigh test), nor were there significantly more direction-tuned neurons that preferred contralateral directions ($p > 0.05$, χ^2 test).

We examined whether individual neurons jointly encoded action and reward outcome by looking for a relationship between tuning sharpness and reward responsiveness. If information about action and reward outcome were encoded independently, there would be no systematic variation in measures of these properties. In Figure 8, the reward responsiveness index is plotted against the tuning sharpness for neurons that were significant for either *RR* or *TS* for either rewarded or unrewarded conditions ($p < 0.05$). We found that postmovement PANs encoded movement and reward information interactively rather than independently. Absolute reward responsiveness and tuning sharpness were negatively correlated (Spearman rank correlation $r_s = -0.43$, $p < 0.01$), indicating that the more sharply tuned a neu-

ron was for saccade direction, the less it discriminated rewarded from unrewarded trials. However, as indicated by the L-shaped distribution of the data in Figure 8, this is not simply a negative linear correlation. Neurons appear to be segregated such that some neurons primarily encoded received outcome, whereas others primarily encoded saccade direction. This conclusion is supported by an analysis of the joint distribution. If reward responsiveness and tuning sharpness were encoded independently across our population, then the joint distribution of these properties should be predicted from their marginal distributions; knowing that a neuron was sharply tuned for direction provides no additional information about its degree of reward responsiveness. We used the marginal medians to predict the joint distribution of $|RR|$ and *TS* under the independence hypothesis. The marginal medians (Fig. 8, dashed lines) predict that 25% of the data points should fall into each of the quadrants defined by their intersection. This independence hypothesis is rejected ($p < 0.01$, χ^2 test). The joint distribution of reward responsiveness and tuning sharpness has fewer than expected sharply tuned neurons with large differential reward responses, which supports the impression given by the paucity of data in the top right quadrant of Figure 8. This result remains true if we adopt a stricter significance level for including neurons in Figure 8 ($p < 0.01$ in-

stead of $p < 0.05$). As expected, a more stringent criterion leads to fewer neurons with tuning sharpness <0.10 being labeled as significantly direction tuned. Importantly however, both the correlation and the marginal median analyses remain highly significant ($r_s = -0.58$, $p < 0.01$ and $p < 0.01$, χ^2 test, respectively). Thus, neurons active during the reward epoch appear to fall into two categories, one that encodes the direction of the just-completed saccade and one that encodes whether or not a reward was received.

In addition, we note that our conclusions were still supported even when we used a different criterion for direction tuning. We selected an alternative criterion using data recorded from PANs that responded primarily before the appearance of the peripheral target (Precue). In our task, because reward expectation was held constant, a Precue neuron could not exhibit direction tuning before the onset of the peripheral target because cue directions were randomly selected on each trial. Thus, Precue neurons serve as a useful control because estimates of the tuning sharpness for these neurons are obtained when the null hypothesis must be true (no direction tuning). As expected, no Precue neurons were deemed significant by our statistical test, and none exhibited a tuning sharpness >0.25 (Fig. 5a). Therefore 0.25 is a justifiable, albeit conservative, alternative criterion for direction tuning. Using this alternative criterion, we find that of the 35 neurons in Figure 8 that were significantly reward responsive, 26 would be categorized as untuned, whereas only 9 would be considered direction tuned. In addition, of the 19 neurons in Figure 8 that were "direction only," 15 remained categorized as direction tuned.
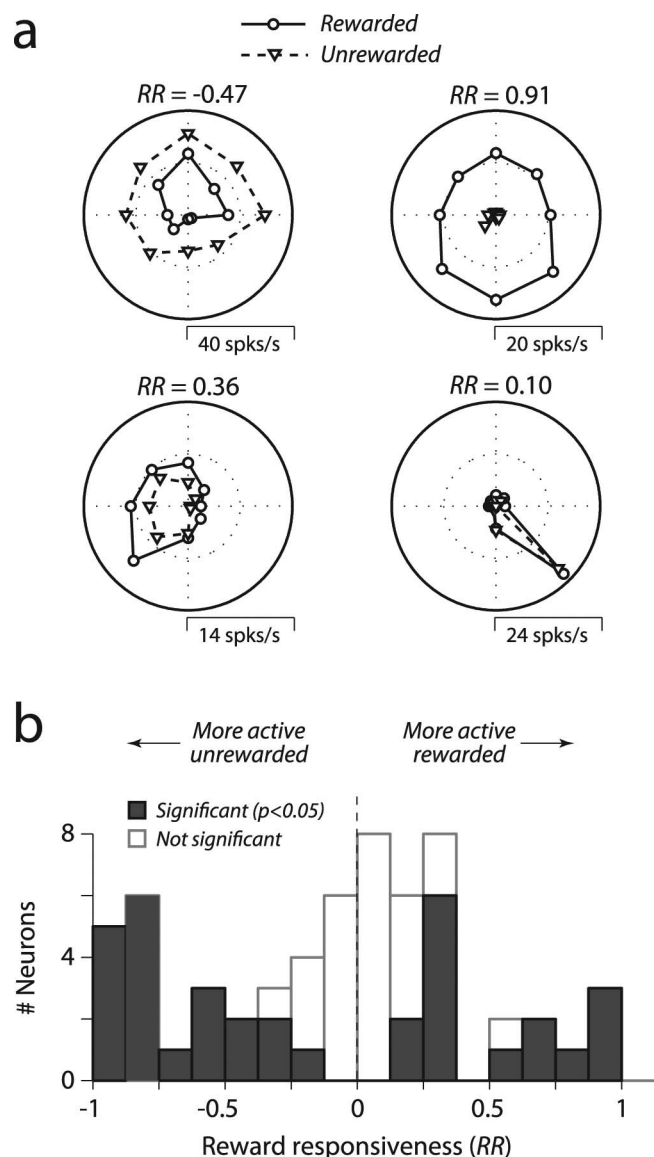
a

b

Figure 6. Neurons with responses during the reward epoch respond differentially to reward delivery. **a**, Example neurons with peak responses after reward onset. The reward responsiveness index for each neuron is listed in the top left corner of each polar plot. **b**, Histogram of *RR*s for all neurons with the first time to half-maximal response after the peak response time half-max windows extending at least 150 ms into the reward epoch (n = 60, 4 Postsaccade neurons, 56 Reward neurons). Only one Postsaccade neuron exhibited a significant *RR* (tuning functions plotted in Fig. 8).

a

b

Figure 7. Neurons responding after saccade execution are direction tuned. **a**, Tuning strength estimated from rewarded plotted against *TS* estimated from unrewarded trials for neurons with a significant tuning strength for either condition (light gray symbols, n = 43, 3 Postsaccade neurons, 40 Reward neurons). The dark filled symbols represent neurons that have jointly significant tuning strengths for rewarded and unrewarded trials (n = 22, 2 Postsaccade neurons, 20 Reward neurons). **b**, Preferred direction estimated from rewarded plotted against preferred direction estimated from unrewarded trials.

Thus, using an alternative criterion for direction selectivity, we continue to find that postmovement neurons segregate with respect to tuning strength and reward responsiveness, consistent with our conclusion that reward-responsive neurons are distinct from direction-tuned neurons.

Together, these results indicate that neurons responsive during the reward epoch encode information about saccade direction as well as received outcome. As was the case for Postsaccade neurons, direction tuning exhibited by these neurons constitutes a retrospective coding of direction. Furthermore, there was relatively little overlap of these properties in individual neurons, and the encoding of direction and received outcome appeared to be performed by largely separate populations of caudate neurons.
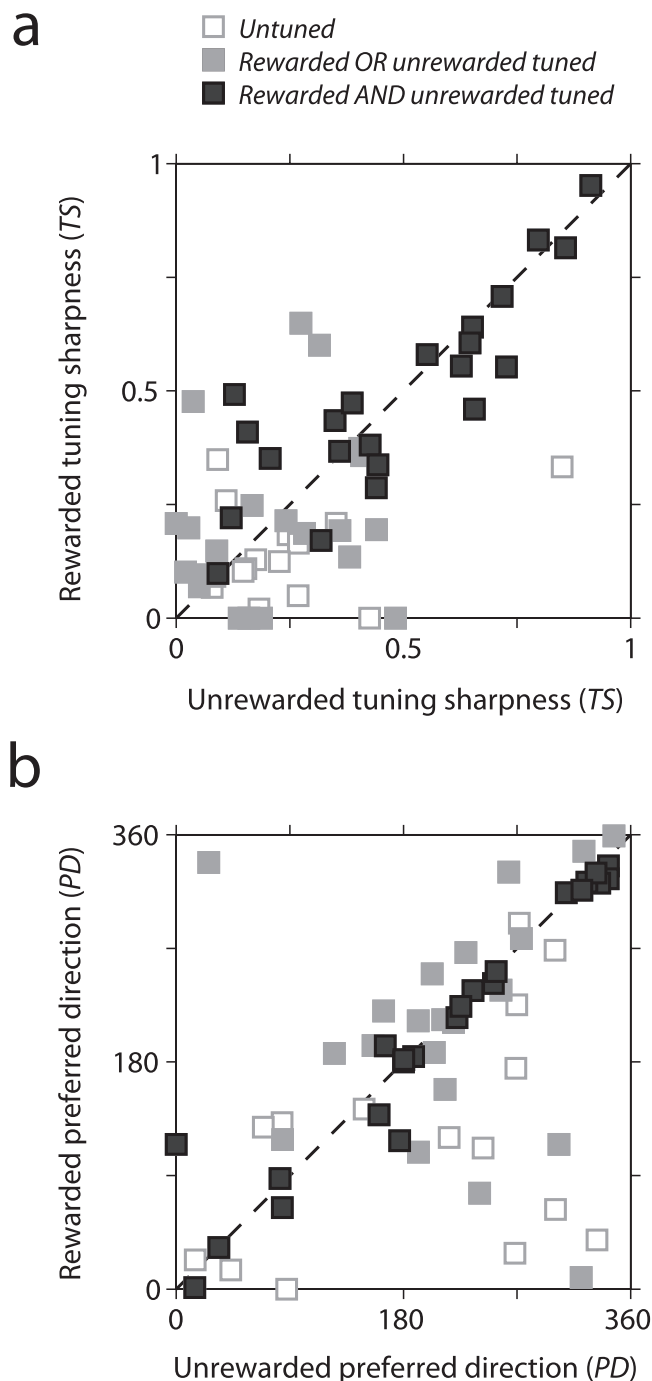
Postmovement responses are aligned to the instructed saccade

It is possible that PAN responses that occurred late in the trial were related to events after the primary, instructed saccade; for example, these responses may be related to the secondary saccade that moves the eye away from the target position after reward delivery. We checked for this possibility by analyzing whether
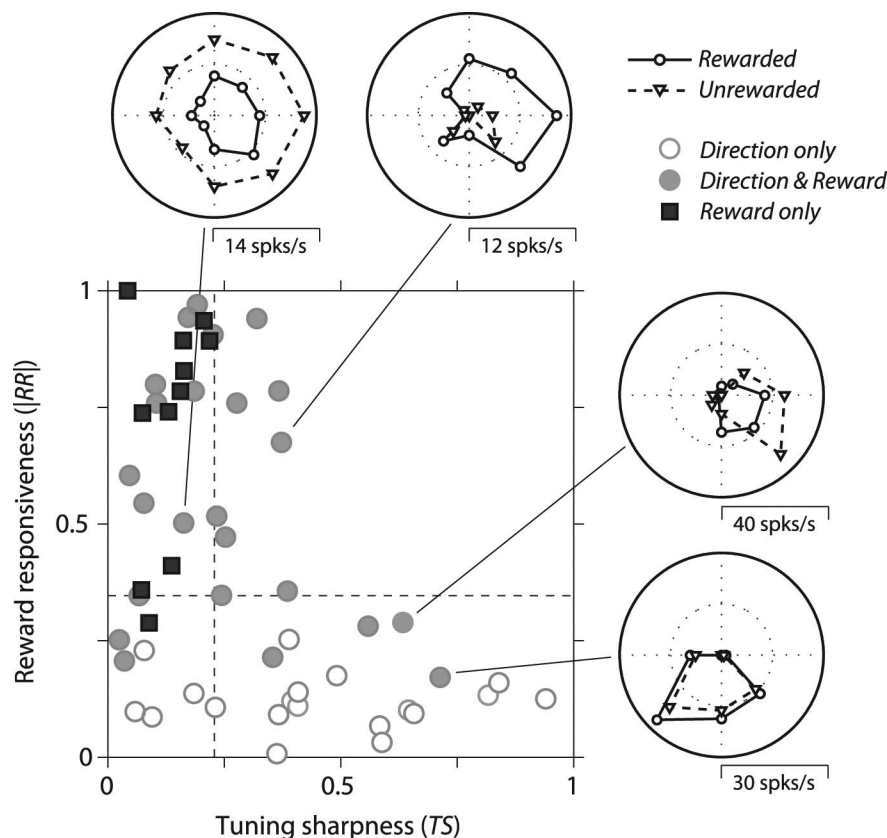
Figure 8. Relation between direction tuning and absolute reward responsiveness for neurons with peak responses after saccade execution. The tuning sharpness along the abscissa is taken as an average of the tuning sharpness for the rewarded and unrewarded conditions, weighted by the reciprocal of the width of the associated confidence interval. Only neurons that are significant for either *RR* or *TS* are plotted (*n* = 54, 3 Postsaccade neurons, 51 Reward neurons). Four example neurons that were jointly significant for both *TS* and *RR* are shown (the single Postsaccade neuron differentially responsive to reward, and 3 Reward neurons). The dashed lines represent the marginal medians. The quadrants defined by the intersection of these lines would each contain 25% of the data points under the hypothesis of independence between tuning sharpness and reward responsiveness.

ture the responses of the small number of neurons that were multiphasic or had slow buildup activity preceding a phasic burst [particularly Saccade neurons (Fig. 3)]. We addressed these aspects of our data by estimating tuning sharpness in 200 ms non-overlapping windows spanning the entire trial for each neuron.

A summary of direction tuning across our population is provided in Figure 10*a*, where the percentage of significant activations is plotted as a function of time. Approximately 30% of our population became significantly tuned for direction shortly after cue presentation, and this fraction remains approximately constant throughout the delay, saccade, and reward epochs, falling to chance levels ~800 ms after reward onset. This percentage is lower than those in the previous sections of this report because it was estimated with respect to the entire population (black line) or all neurons responding above baseline for each window (dashed black line). Because PAN response profiles were distributed throughout the trial, the number of neurons responding vigorously in any particular window was relatively small. Despite this, it appears that an approximately constant number of neurons encode cue or saccade direction for a significant fraction of a trial.

How does the percentage of significantly tuned neurons relate to response sharpness? The mean population response (Fig. 10*a*, blue line) closely matches the percentage of significantly tuned neurons (Fig. 10*a*, black lines) except for epochs preceding cue onset, where the functions deviate as a result of untuned Precue neurons being included in the average. Because our tuning sharpness measure (*TS*) does not depend on our response strength measure (*M*), the correspondence between significance and average response in Figure 10*a* does not obviously reveal how tuning sharpness evolves through a trial. It is possible that tuning sharpness depends on response strength; different neurons underlying the average response come to represent cue or saccade direction at different times throughout the trial. For example, Cue/Delay neurons could be primarily tuned immediately after cue onset when their firing rates were highest, whereas Saccade neurons could be primarily tuned immediately preceding the saccade when their firing rates were highest. This is suggested by the temporal tiling of responses illustrated in Figure 3, as well as by Figure 10*b*, which shows the average response separated by neuron category. Alternatively, tuning sharpness might not depend on response strength, and sharp tuning for direction might exist despite large changes in firing rate. The average tuning sharpness (Fig. 10*c*) supports the latter alternative, revealing a dissociation of tuning sharpness from peak response time and average response. The Saccade and Postsaccade populations, in particular, show average tuning sharpnesses that do not match their respective average response strengths. In summary, these data indicate that the sharpest tuning is not aligned to the maximal response. Rather, these populations become tuned shortly after cue presen-

individual neuronal responses were better aligned to the primary or secondary saccade. Figure 9*a* plots the responses of two post-movement PANs, aligned to the onset of both the primary and secondary saccades. In both cases, aligning to the secondary saccade reduces the peak response and distributes the response more widely in time. We used the peak response and the half-width, the time between the two points of half the peak response, to quantify how "phasic" a response was when aligned to the primary or secondary saccade. The data for individual postmovement neurons is plotted in Figure 9*b*, which shows that the great majority of postmovement neurons exhibited responses that were better aligned to the primary rather than secondary saccade. Figure 9*c* shows that similar results are obtained when the same analysis was performed on Saccade neurons. These results support our conclusion that postmovement neurons convey information about events temporally linked to the primary, instructed saccade.

Dynamic tuning analysis

We restricted analyses in the preceding sections to a single temporal window containing the peak response for each neuron. Although our categorization (Precue, Cue/Delay, Saccade, etc.) provides information regarding how tuning sharpness evolves through a trial, this provides fairly coarse information because each category included individual neurons responding at different times. Furthermore, the half-max windows did not fully cap-
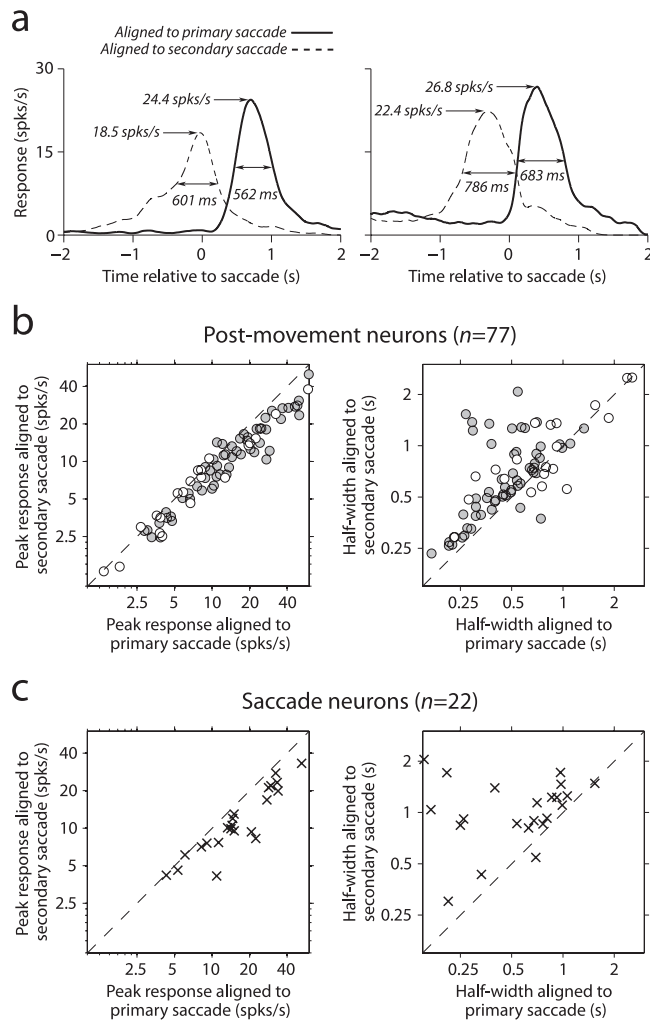
Figure 9. Postmovement responses align better to the primary instructed saccade than to the secondary saccade. ***a***, Example postmovement neurons, showing the metrics used to quantify differences between the primary saccade to the instructed target and the secondary saccade after reward delivery. ***b***, Summary for postmovement neurons (21 Postsaccade, open symbols; 56 Reward, shaded symbols). The response is higher and narrower for the majority of neurons. ***c***, Neurons most active immediately preceding the primary saccade are plotted for comparison.

Figure 10. Dynamic tuning analysis, population averages. ***a***, Population significance of *TS* as a function of time. The solid black line represents the percentage significance from the entire population (*n* = 149). The dashed line shows percentage significance from all neurons responding above baseline in each window. The mean population response (blue line) was estimated by averaging response strength (*M*) across neurons for each 200 ms window. ***b***, Average response (*M*) for each neuron category. ***c***, Average *TS* for each neuron category. Data are plotted for those windows in which at least 15 neurons responded at least 1.5 times above baseline.

tation, and remain tuned at an approximately constant tuning sharpness until reward onset.

The tuning sharpness calculated using multiple smaller windows agrees with that computed using the half-max window for each neuron (Fig. 5), suggesting that those neurons with prelude or buildup activity preceding the peak response should exhibit direction tuning early in the trial. This is verified by examining the tuning sharpness for individual neurons (Fig. 11). There are subsets of neurons from the Cue/Delay, Saccade, and Postsaccade categories that become tuned shortly after cue onset [neurons in Fig. 11 are ordered according to peak times (Fig. 3)]. Some of these neurons remain tuned for much of the time between cue onset and saccade execution (horizontal streaks across Fig. 11 with approximately the same gray level indicate approximately constant tuning sharpness). Moreover, comparing Figures 3 and 11 reveals that tuning sharpness is not necessarily related to overall response strength; some neurons show significant direction tuning well outside of the half-max windows.

Reward neurons were quite different. These neurons rarely showed direction tuning early in the trial, but instead showed
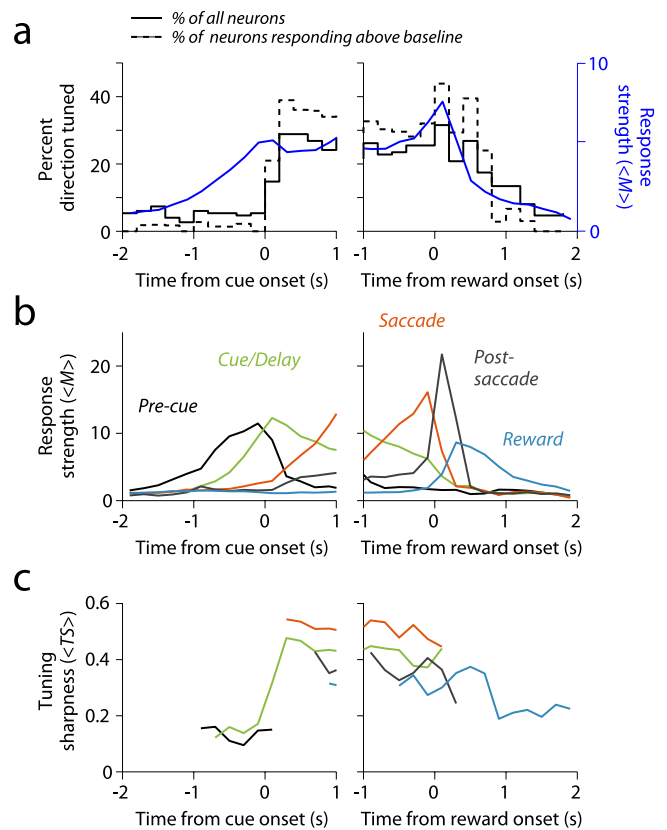
direction tuning primarily restricted to the reward epoch. This is less surprising, however, because Reward neurons showed very little response (Fig. 10*b*) until around the expected time of reward delivery. The corresponding absolute reward responsiveness for these neurons is plotted on the bottom right side of Figure 11*a*. For those neurons that have significant tuning for at least one window during the reward epoch (*n* = 53), there is a significant negative correlation between |*RR*| and the maximum *TS* (r_s = −0.62, *p* < 0.001). Resorting the data for the subset of neurons with significant *RR* according to the absolute magnitude of *RR* (Fig. 11*b*) confirms that neurons with large reward responsiveness show little or no direction tuning, whereas most of the sharply tuned neurons show little reward responsiveness. The neurons with the sharpest direction tuning during the reward epoch are not plotted in Figure 11*b* because their *RR*s were not significantly different from zero; they do not encode received outcome. These results are consistent with our previous analyses based on tailored temporal analysis windows (Fig. 8).

Recording locations

We targeted our recordings to the body of the caudate nucleus based on previous work by Hikosaka et al. (1989a), who showed that this portion of the striatum contains neurons related to saccadic eye movements. Figure 12, *a* and *b*, shows our recording locations. Figure 12*c* illustrates the spatial distribution of the PANs we recorded, aligned to the anterior commissure. We did
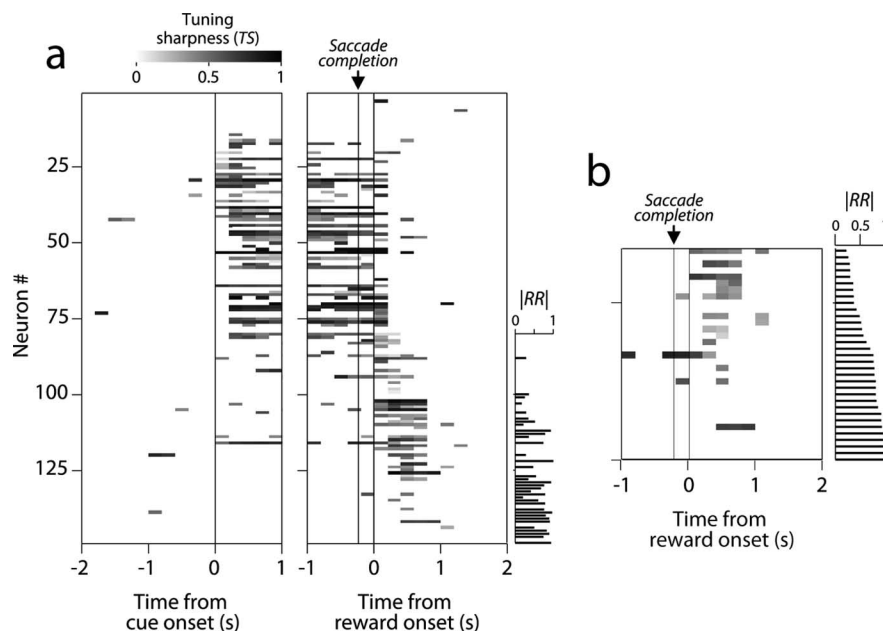
Figure 11. Dynamic tuning analysis, individual data. *a*, *TS* for each neuron plotted as a function of time. The order of neurons is the same as in Figure 3. *TS* is only plotted for those windows in which direction tuning is significant ($p < 0.05$), and the neuronal response was above baseline. *b*, Neurons with significant *RR*, resorted according to increasing |*RR*|.

not observe any obvious clustering of neurons according to the time of peak response; neurons with very different peak responses could be found intermixed throughout the region of the caudate nucleus we recorded from.

Discussion

Hikosaka and colleagues found that some caudate neurons respond selectively for eye movements early enough to initiate saccades by disinhibiting the superior colliculus (Hikosaka et al., 1993). They also found that some caudate neurons respond in a manner not directly related to movement initiation (Hikosaka et al., 1989c). Postmovement activity is also common in the striatum and may play a role in evaluative processing after goal-directed movements (Tremblay et al., 1998; Cromwell and Schultz, 2003; Fujii and Graybiel, 2005; Williams and Eskandar, 2006). Clarifying how postmovement neurons encode movement properties is critical for understanding both the types of information the striatum transmits to downstream brain structures and the role of this structure in reinforcement learning.

We found that actions and outcomes were encoded by largely separate populations of postmovement PANs in a probabilistically rewarded saccade task. One subset of neurons was tuned for the direction of the immediately preceding saccade. Another subset responded differentially on rewarded versus unrewarded trials and was relatively untuned for saccade direction. We also compared the response properties of postmovement PANs with those of premovement PANs, providing a quantitative characterization of these neurons throughout the planning, execution, and feedback stages of instructed eye movements.

Retrospective direction tuning

Postmovement neurons exhibited slightly lower tuning sharpness than premovement neurons. Although some postmovement neurons had prelude activity tuned for the direction of the impending saccade, many neurons did not show tuning until later, often during the reward epoch and sometimes starting after the typical offset times of burst neurons in the superior colliculus that

generate saccades (e.g., Sparks, 1978). It has previously been shown that many pallidal neurons respond too late to initiate arm movements (Anderson and Horak, 1985), and one hypothesis is that this activity influences muscle activity after movement initiation. The ballistic nature of saccades makes this explanation unlikely in our results. The postmovement activity we observed also appears too late and variable to be obviously useful as a corollary discharge (Wurtz and Sommer, 2004), although it may play a role in organizing sequences of movements (Graybiel, 1995). Postmovement activity may function to signal different task components such as the completion and accuracy of the instructed action, the termination of fixation, or the impending reward uncertainty (Fujii and Graybiel, 2003).

Another possibility is that postmovement activity provides feedback about the consequences of recently executed actions. Reinforcement learning models commonly incorporate a short-term memory, termed an eligibility trace, that keeps recently executed actions eligible for plasticity (Sutton and Barto, 1998). Eligibility traces help associate reinforcers with specific actions and can increase learning speed, particularly in situations in which multiple events intervene between actions and outcomes. It is thus possible that retrospective direction coding in the caudate and the pallidum may be related to the eligibility traces of learning theory. This is consistent with the hypothesis that activity in the basal ganglia promotes learning in cortical areas (Wise et al., 1996) via projections through the substantia nigra pars reticulata (SNr) and thalamus (Ilinsky et al., 1985; Middleton and Strick, 1994), an idea supported by recent experiments (Fujii and Graybiel, 2005; Pasupathy and Miller, 2005).

The direction tuning of postmovement neurons was broad in two senses; there were individual neurons with broad tuning, and tuning sharpness varied widely across the neuronal population. Some PANs exhibited statistically significant, but weak, direction tuning. Because it is unclear how the activity of caudate neurons is decoded in downstream brain structures, it remains possible that weakly direction-tuned neurons participate in encoding movement direction. After all, neurons in the SNr are broadly tuned (Hikosaka and Wurtz, 1983), yet projections from these neurons to the superior colliculus are important for initiating precise saccades (Hikosaka and Wurtz, 1985).

We also observed that tuning sharpness varied widely across our population, consistent with earlier observations that the sizes of movement fields of caudate neurons are variable (Hikosaka et al., 1989a). This does not pose a difficulty insofar as it is possible to decode precise information from a population of neurons with heterogeneous direction tuning (Jazayeri and Movshon, 2006), although it is an open question whether downstream brain structures can appropriately decode this information.

Reward responses

A separate population of postmovement PANs encoded whether or not a reward had been received. Our work extends previous descriptions of reward-related responses in the dorsal striatum,

which have only examined postmovement responses when the outcomes are fully expected (Apicella et al., 1991; Hollerman et al., 1998; Cromwell and Schultz, 2003) or during learning when reward expectations are changing (Tremblay et al., 1998; Williams and Eskandar, 2006). In our task, rewards were delivered with a fixed probability, allowing us to dissociate reward outcome from correct task performance and changing reward expectations.

The prevalence of reward responses is consistent with functional magnetic resonance imaging studies showing increased activity in the dorsal striatum in response to rewards (Delgado et al., 2000; Knutson et al., 2001; Haruno et al., 2004) and reward prediction errors (RPEs) (O'Doherty et al., 2004; Haruno and Kawato, 2006). RPEs, together with eligibility traces, are core elements in temporal-difference models of reinforcement learning (Sutton and Barto, 1998). It is possible that the reward-responsive neurons we observed encode RPEs, which would be consistent with projections to the striatum from midbrain dopamine neurons, which encode RPEs (Schultz, 1998; Bayer and Glimcher, 2005).

Segregation of action and outcome
Postmovement PANs encoded information about actions and obtained outcomes interactively; the fact that a neuron was strongly direction tuned predicted that it was unlikely to also be reward responsive, suggesting that movement direction and obtained outcomes are encoded by two distinct groups of neurons at the level of the caudate nucleus (cf. Schmitzer-Torbert and Redish, 2004).

How information flows through the basal ganglia is the subject of much inquiry (Alexander et al., 1986; Bar-Gad et al., 2003). Our observations in postmovement neurons suggest that at the level of the striatum, information about actions and received outcomes is segregated, although it remains possible that these neurons also encode other variables (e.g., reward expectation, discussed below). In addition, information might be specifically integrated downstream of the striatum. Current evidence suggests that this information is combined independently in pallidal neurons (Arkadir et al., 2004) and that the fraction of neurons encoding movement direction increases as information is transmitted from the putamen to the pallidum (Pasquereau et al., 2007).

Relation to work manipulating reward expectation
Our findings relate to work on the relationship between action and reward expectation. Hikosaka and colleagues (Kawagoe et al., 1998) examined the encoding of movement direction and reward expectation in caudate PANs using a task in which monkeys were rewarded for saccades to one of four targets, but were required to make saccades to all four. They showed that reward expectation can modulate the direction tuning of premovement PANs; the direction tuning of some neurons reflects the direction
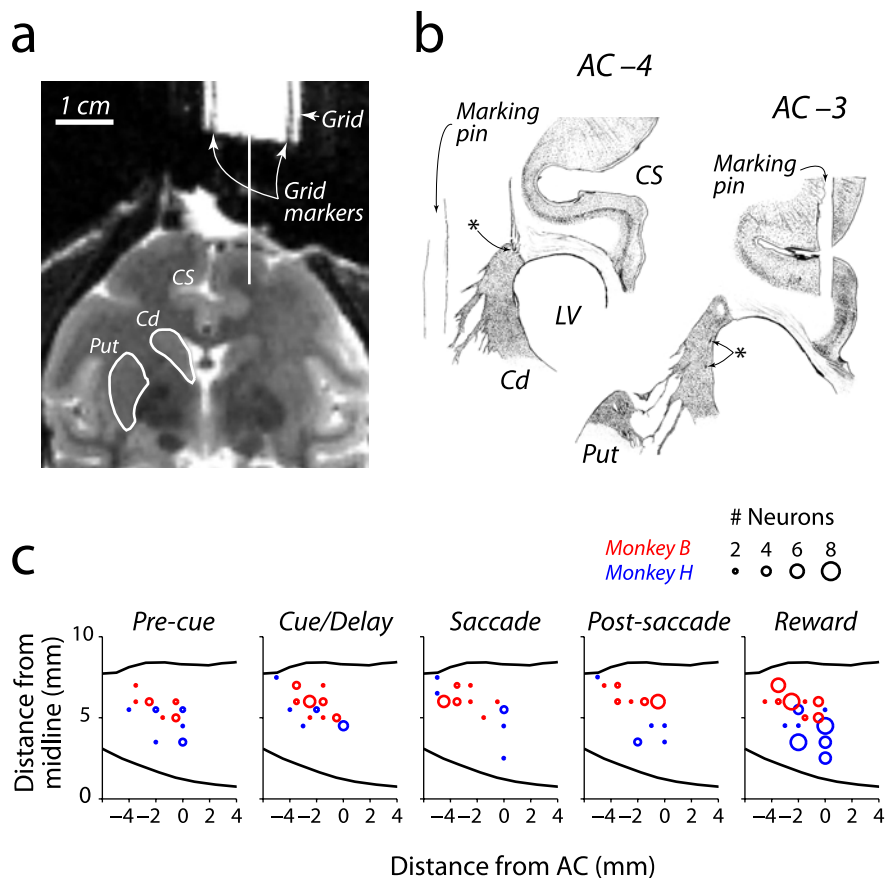

Figure 12. Recording locations. **a**, Structural MRI image for monkey B. We used vitamin E-filled capillary tubes as grid markers, which allowed us to section the image in the plane of the recording grid. The white line through the center of the grid represents the angle of approach, and its length is the same as the guide tube used during recording. Cd, Caudate nucleus; CS, cingulate sulcus; Put, putamen. **b**, Camera lucida drawings of histological sections for monkey H. The position of each section is estimated relative to the anterior commissure (AC). The asterisk in each section indicates the location of electrolytic lesions recovered in the caudate nucleus. LV, Lateral ventricle. **c**, Recording locations for task-related PANs, plotted separately for each neuronal category. The caudate outline represents the boundaries of the nucleus when viewed from above as taken from the atlas of Francois et al. (1996).

of the rewarded target, consistent with the idea that these neurons interactively encode movement direction and reward expectation (Kawagoe et al., 1998). In some PANs, reward expectation can sharpen direction tuning. The degree of sharpening has been measured in Precue PANs responding before target onset, and the direction tuning induced by reward expectation in these neurons is comparable in tuning sharpness to what we observed in postmovement neurons (cf. Takikawa et al., 2002). In addition to these neurons, Hikosaka and colleagues have also identified PANs that respond for all movements in a block when reward is associated with a specific target (Ding and Hikosaka, 2006; Kobayashi et al., 2007). Moreover, premovement PANs can be modulated by relative expected value (Cromwell and Schultz, 2005; Samejima et al., 2005). These results indicate that premovement PANs can be powerfully modulated by reward expectation.

Postmovement PANs may also be modulated by reward expectation, a possibility our results do not exclude. Ding and Hikosaka (2006) showed that postsaccadic PAN activity can be modulated by reward expectation, although less frequently than premovement PAN activity. They suggest that the caudate contributes to evaluating rewards associated with particular actions, a hypothesis elaborated by our findings. Related work in premovement neurons (Kawagoe et al., 1998; Pasquereau et al.,

2007) suggests that reward expectation may also modulate direction tuning in postmovement neurons. This type of postmovement activity might function to convey reward expectations to midbrain dopamine neurons via known projections (Joel and Weiner, 2000) to enable the computation of action-specific reward prediction errors (Houk et al., 1995b; Doya, 2000). Another possibility is that postmovement activity reactivates their target neurons to coincide with the dopamine efflux that occurs in response to unpredicted rewards, thereby promoting plasticity in downstream structures that is specific to rewarded actions (functioning as an eligibility trace).

Conclusion

We characterized saccadic direction tuning in striatal PANs active during a probabilistically rewarded saccade task. Many neurons had peak responses after saccade execution. These postmovement PANs were functionally segregated, and may play different roles in evaluative updating; one population encodes information about the saccade just executed, whereas another encodes the received outcome. We also examined how direction tuning evolved through time across PANs, and found that selectivity for saccade direction was sustained after cue onset to well after saccade execution. Together, these results suggest that a subset of caudate PANs may participate in evaluating the outcomes of actions.

References

Alexander GE, DeLong MR, Strick PL (1986) Parallel organization of functionally segregated circuits linking basal ganglia and cortex. Annu Rev Neurosci 9:357–381.

Anderson ME, Horak FB (1985) Influence of the globus pallidus on arm movements in monkeys. III. Timing of movement-related information. J Neurophysiol 54:433–448.

Aosaki T, Tsubokawa H, Ishida A, Watanabe K, Graybiel AM, Kimura M (1994) Responses of tonically active neurons in the primate's striatum undergo systematic changes during behavioral sensorimotor conditioning. J Neurosci 14:3969–3984.

Apicella P, Ljungberg T, Scarnati E, Schultz W (1991) Responses to reward in monkey dorsal and ventral striatum. Exp Brain Res 85:491–500.

Apicella P, Scarnati E, Ljungberg T, Schultz W (1992) Neuronal activity in monkey striatum related to the expectation of predictable environmental events. J Neurophysiol 68:945–960.

Apicella P, Legallet E, Trouche E (1997) Responses of tonically discharging neurons in the monkey striatum to primary rewards delivered during different behavioral states. Exp Brain Res 116:456–466.

Arkadir D, Morris G, Vaadia E, Bergman H (2004) Independent coding of movement direction and reward prediction by single pallidal neurons. J Neurosci 24:10047–10056.

Bar-Gad I, Morris G, Bergman H (2003) Information processing, dimensionality reduction and reinforcement learning in the basal ganglia. Prog Neurobiol 71:439–473.

Barnes TD, Kubota Y, Hu D, Jin DZ, Graybiel AM (2005) Activity of striatal neurons reflects dynamic encoding and recoding of procedural memories. Nature 437:1158–1161.

Bayer H, Glimcher P (2005) Midbrain dopamine neurons encode a quantitative reward prediction error signal. Neuron 47:129–141.

Cromwell HC, Schultz W (2003) Effects of expectations for different reward magnitudes on neuronal activity in primate striatum. J Neurophysiol 89:2823–2838.

Cromwell HC, Schultz W (2005) Relative reward processing in primate striatum. Exp Brain Res 162:520–525.

Delgado MR, Nystrom LE, Fissell C, Noll DC, Fiez JA (2000) Tracking the hemodynamic responses to reward and punishment in the striatum. J Neurophysiol 84:3072–3077.

Ding L, Hikosaka O (2006) Comparison of reward modulation in the frontal eye field and caudate of the macaque. J Neurosci 26:6695–6703.

Doya K (2000) Complementary roles of basal ganglia and cerebellum in learning and motor control. Curr Opin Neurobiol 10:732–739.

Efron B, Tibshirani R (1994) An introduction to the bootstrap. Boca Raton, FL: Chapman Hall/CRC.

Fisher NI (1993) Statistical analysis of circular data. Cambridge, UK: Cambridge UP.

Francois C, Yelnik J, Percheron G (1996) A stereotaxic atlas of the basal ganglia in macaques. Brain Res Bull 41:151–158.

Fuchs AF, Robinson DA (1966) A method for measuring horizontal and vertical eye movement chronically in the monkey. J Appl Physiol 21:1068–1070.

Fujii N, Graybiel AM (2003) Representation of action sequence boundaries by macaque prefrontal cortical neurons. Science 301:1246–1249.

Fujii N, Graybiel AM (2005) Time-varying covariance of neural activities recorded in striatum and frontal cortex as monkeys perform sequential-saccade tasks. Proc Natl Acad Sci USA 102:9032–9037.

Glimcher PW, Ciaramitaro VM, Platt ML, Bayer HM, Brown MA, Handel A (2001) Application of neurosonography to experimental physiology. J Neurosci Methods 108:131–144.

Graybiel AM (1995) Building action repertoires: memory and learning functions of the basal ganglia. Curr Opin Neurobiol 5:733–741.

Gribble PL, Scott SH (2002) Method for assessing directional characteristics of non-uniformly sampled neural activity. J Neurosci Methods 113:187–197.

Hanes DP, Thompson KG, Schall JD (1995) Relationship of presaccadic activity in frontal eye field and supplementary eye field to saccade initiation in macaque: Poisson spike train analysis. Exp Brain Res 103:85–96.

Haruno M, Kawato M (2006) Different neural correlates of reward expectation and reward expectation error in the putamen and caudate nucleus during stimulus-action-reward association learning. J Neurophysiol 95:948–959.

Haruno M, Kuroda T, Doya K, Toyama K, Kimura M, Samejima K, Imamizu H, Kawato M (2004) A neural correlate of reward-based behavioral learning in caudate nucleus: a functional magnetic resonance imaging study of a stochastic decision task. J Neurosci 24:1660–1665.

Hikosaka O, Wurtz RH (1983) Visual and oculomotor functions of monkey substantia nigra pars reticulata. III. Memory-contingent and saccade responses. J Neurophysiol 49:1268–1284.

Hikosaka O, Wurtz RH (1985) Modification of saccadic eye movements by GABA-related substances. II. Effects of muscimol in monkey substantia nigra pars reticulata. J Neurophysiol 53:292–308.

Hikosaka O, Sakamoto M, Usui S (1989a) Functional properties of monkey caudate neurons. I. Activities related to saccadic eye movements. J Neurophysiol 61:780–798.

Hikosaka O, Sakamoto M, Usui S (1989b) Functional properties of monkey caudate neurons. II. Visual and auditory responses. J Neurophysiol 61:799–813.

Hikosaka O, Sakamoto M, Usui S (1989c) Functional properties of monkey caudate neurons. III. Activities related to expectation of target and reward. J Neurophysiol 61:814–832.

Hikosaka O, Sakamoto M, Miyashita N (1993) Effects of caudate nucleus stimulation on substantia nigra cell activity in monkey. Exp Brain Res 95:457–472.

Hollerman JR, Tremblay L, Schultz W (1998) Influence of reward expectation on behavior related neuronal activity in primate striatum. J Neurophysiol 80:947–963.

Houk JC, Davis JL, Beiser DG, eds (1995a) Models of information processing in the basal ganglia. Cambridge, MA: MIT.

Houk JC, Adams JL, Barto AG (1995b) A model of how the basal ganglia generate and use neural signals that predict reinforcement. In: Models of information processing in the basal ganglia (Houk JC, Davis JL, Beiser DG, eds). Cambridge, MA: MIT.

Ilinsky IA, Jouandet ML, Goldman-Rakic PS (1985) Organization of the nigrothalamocortical system in the rhesus monkey. J Comp Neurol 236:315–330.

Itoh H, Nakahara H, Hikosaka O, Kawagoe R, Takikawa Y, Aihara K (2003) Correlation of primate caudate neural activity and saccade parameters in reward-oriented behavior. J Neurophysiol 89:1774–1783.

Jazayeri M, Movshon JA (2006) Optimal representation of sensory information by neuronal populations. Nat Neurosci 9:690–696.

Joel D, Weiner I (2000) The connections of the dopaminergic system with the striatum in rats and primates: an analysis with respect to the functional and compartmental organization of the striatum. Neuroscience 96:451–474.

Judge SJ, Richmond BJ, Chu FC (1980) Implantation of magnetic search coils for measurement of eye position: an improved method. Vision Res 20:535–538.

Kawagoe R, Takikawa Y, Hikosaka O (1998) Expectation of reward modulates cognitive signals in the basal ganglia. Nat Neurosci 1:411–416.

Kimura M, Rajkowski J, Evarts E (1984) Tonically discharging putamen neurons exhibit set-dependent responses. Proc Natl Acad Sci USA 81:4998–5001.

Kobayashi S, Kawagoe R, Takikawa Y, Koizumi M, Sakagami M, Hikosaka O (2007) Functional differences between macaque prefrontal cortex and caudate nucleus during eye movements with and without reward. Exp Brain Res 176:341–355.

Knutson B, Fong GW, Adams CM, Varner JL, Hommer D (2001) Dissociation of reward anticipation and outcome with event-related fMRI. Neuroreport 12:3683–3687.

Lauwereyns J, Takikawa Y, Kawagoe R, Kobayashi S, Koizumi M, Coe B, Sakagami M, Hikosaka O (2002a) Feature-based anticipation of cues that predict reward in monkey caudate nucleus. Neuron 33:463–473.

Lauwereyns J, Watanabe K, Coe B, Hikosaka O (2002b) A neural correlate of response bias in monkey caudate nucleus. Nature 418:413–417.

Lee IH, Seitz AR, Assad JA (2006) Activity of tonically active neurons in the monkey putamen during initiation and withholding of movement. J Neurophysiol 95:2391–2403.

Middleton FA, Strick PL (1994) Anatomical evidence for cerebellar and basal ganglia involvement in higher cognitive function. Science 266:458–461.

O'Doherty J, Dayan P, Schultz J, Deichmann R, Friston K, Dolan RJ (2004) Dissociable roles of ventral and dorsal striatum in instrumental conditioning. Science 304:452–454.

Pasquereau B, Nadjar A, Arkadir D, Bezard E, Goillandeau M, Bioulac B, Gross CE, Boraud T (2007) Shaping of motor responses by incentive values through the basal ganglia. J Neurosci 27:1176–1183.

Pasupathy A, Miller EK (2005) Different time courses of learning-related activity in the prefrontal cortex and striatum. Nature 433:873–876.

Paulin MG, Hoffman LF (2001) Optimal firing rate estimation. Neural Netw 14:877–881.

Platt ML, Glimcher PW (1997) Responses of intraparietal neurons to saccadic targets and visual distractors. J Neurophysiol 78:1574–1589.

Rolls ET, Thorpe SJ, Maddison SP (1983) Responses of striatal neurons in the behaving monkey. 1. Head of the caudate nucleus. Behav Brain Res 7:179–210.

Samejima K, Ueda Y, Doya K, Kimura M (2005) Representation of action-specific reward values in the striatum. Science 310:1337–1340.

Schmitzer-Torbert N, Redish AD (2004) Neuronal activity in the rodent dorsal striatum in sequential navigation: separation of spatial and reward responses on the multiple t task. J Neurophysiol 91:2259–2272.

Schultz W (1998) Predictive reward signal of dopamine neurons. J Neurophysiol 80:1–27.

Schultz W, Apicella P, Romo R, Scarnati E (1995) Context-dependent activity in the primate striatum reflecting past and future behavioral events. In: Models of information processing in the basal ganglia (Houk JC, Davis JL, Beiser DG, eds), pp 11–28. Cambridge, MA: MIT.

Sparks DL (1978) Functional properties of neurons in the monkey superior colliculus: coupling of neuronal activity and saccade onset. Brain Res 156:1–16.

Stark E, Abeles M (2005) Applying resampling methods to neurophysiological data. J Neurosci Methods 145:133–144.

Sutton RS, Barto AG (1998) Reinforcement learning: an introduction. Cambridge, MA: MIT.

Takikawa Y, Kawagoe R, Hikosaka O (2002) Reward-dependent spatial selectivity of anticipatory activity in monkey caudate neurons. J Neurophysiol 87:508–515.

Tremblay L, Hollerman JR, Schultz W (1998) Modifications of reward expectation-related neuronal activity during learning in primate striatum. J Neurophysiol 80:964–977.

Watanabe K, Hikosaka O (2005) Immediate changes in anticipatory activity of caudate neurons associated with reversal of position-reward contingency. J Neurophysiol 94:1879–1887.

Watanabe K, Lauwereyns J, Hikosaka O (2003) Neural correlates of rewarded and unrewarded eye movements in the primate caudate nucleus. J Neurosci 23:10052–10057.

Williams ZM, Eskandar EN (2006) Selective enhancement of associative learning by microstimulation of the anterior caudate. Nat Neurosci 9:562–568.

Wise SP, Murray EA, Gerfen CR (1996) The frontal cortex-basal ganglia system in primates. Crit Rev Neurobiol 10:317–356.

Wurtz RH, Sommer MA (2004) Identifying corollary discharges for movement in the primate brain. Prog Brain Res 144:47–60.